Arctera Processing and Analysis Module

Overview

The Arctera Processing and Analysis Module enables rapid and accurate filtering, processing, searching, and data analysis in multiple formats and languages. Using the Arctera™ eDiscovery Platform, corporations, government agencies, and law firms perform early case assessments and rapidly cull down data, thereby reducing overall electronic discovery costs. As an integrated part of the eDiscovery Platform, the Processing and Analysis Module supports the iterative workflows required during real-world electronic discovery. This solution delivers deep insight into case facts and enables a new level of transparency and defensibility throughout the electronic discovery process.

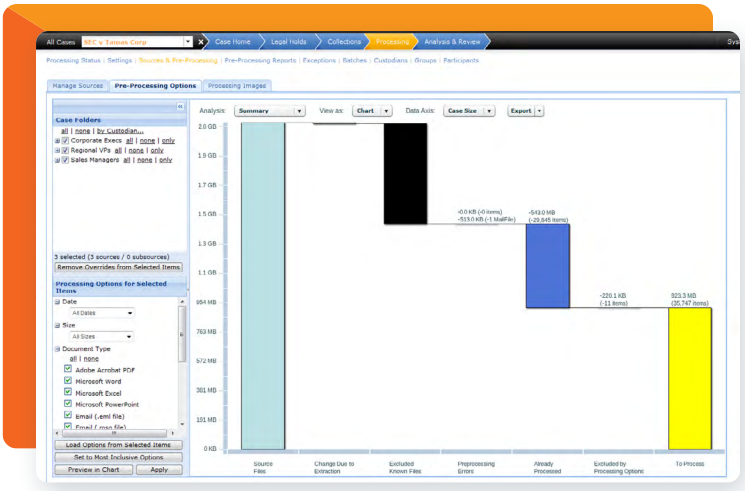

Figure 1. Graphical summary of downstream processing.

Processing

Advanced pre-processing filters—Significantly reduce downstream processing and review costs. This solution enables users to filter data by custodian, date, strong file type, and file size prior to processing. The eDiscovery Platform also provides one-click filtering of custom file and "NIST list" items.

Pre-processing analytics—Rapidly confirm that all case data has been collected and allow for accurate estimation of eDiscovery budgets and timelines. The solution quickly summarizes overall document set characteristics and presents detailed analysis by custodian, timeline, and file type.

Rapid processing—Behind the scenes processing simplifies consolidation for administrators. Concurrent processing and review enables continuous review functionality while processing new documents in the same case.

Robust file support—Processes and analyzes over 400 different file types from data stores and network file shares including Microsoft Office® and PDF documents, various email formats such as PST, NSF, MBOX, OST, and EMLX, attachments, and Guidance Software® LEF and EO1 files. Translates text in image files to searchable content with integrated OCR and identifies over 40 types of hidden content.

Multi-language support—Provides full Unicode compliance and supports English, Western European, Eastern European, Cyrillic, and Asian languages. The eDiscovery Platform also enables automatic language identification of documents and provides exact document counts by language type across the entire data set.

Analysis

Find similar—Concurrently reviewing documents with similar content accelerates the review process and ensures greater tag consistency. Using a dynamically configurable similarity threshold, reviewers are able to easily identify and view emails, attachments, and loose files with similar content to the document under review.

Classification—The information classifier performs a full text index of all data it is analyzing, and stores this information in an internal database. During this analysis it searches for text that matches the conditions in the classification policies. If it finds a match, it adds the appropriate classification tag to the information it stores about the data. You can use the index of text, classification tags and metadata to search within the data to identify all items that apply to the investigation. You can then use this information for early case assessment, which provides easy filtering of the collection, ensuring a streamlined list of documents for review.

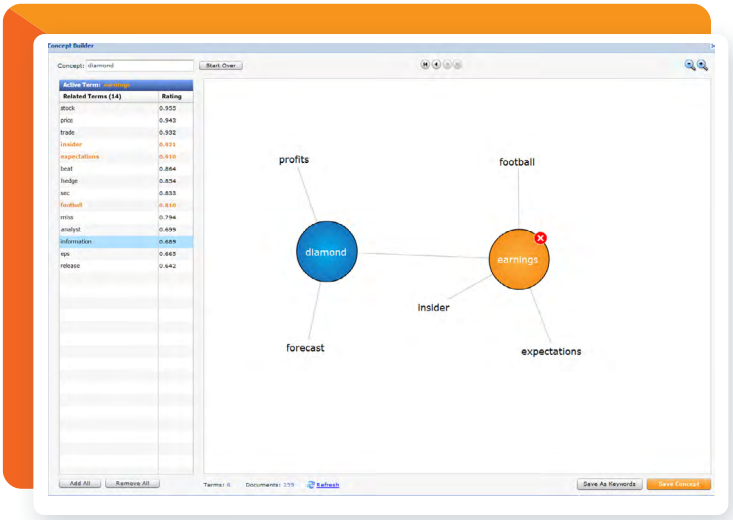

Figure 2. Concept search explorer: Provides a visual interface to dynamically explore and discover new, relevant concepts.

People analytics—Analyze individual and group-to-group communications within a company, to customers, suppliers, and partners. Users can easily access a list of top custodians for a search or monitor communications between divisions.

Term analytics—Leveraging natural language algorithms to analyze noun phrases, users can uncover secret project names and code words that may be relevant to a case or investigation.

Near-duplicate identification—Easily identify, view, and tag near-duplicate emails, attachments, and loose files. Differences in near duplicate documents are automatically highlighted for comparison.

Search

Transparent keyword search—Enables a more defensible and collaborative electronic discovery search process and enhances the ability to cull irrelevant information.

Keyword search preview—Provides matching keyword variations prior to running a search. Users can selectively include relevant variations or exclude false positive variations.

Keyword search filters—Enables real-time search result filtering for individual queries or variations and allows users to sample the filtered documents.

Keyword search report—Provides comprehensive reporting that documents all search criteria and provides detailed analytics of the results.

Multi-keyword search—Has the ability to run up to 100 queries simultaneously, dramatically decreasing the time needed to evaluate the effectiveness of keyword searches.

Advanced search—Has the ability to construct advanced searches based on numerous metadata and derived fields. It supports both stemmed and unstemmed (literal) searches and provides power-user capabilities including Boolean, wildcard, fuzzy, nested proximity searches, and prediction scores.

Auto-filters—Automatically groups search results by metadata fields such as tag, sender domain, document type, custodian, language type, and prediction score and displays exact hit counts across the entire search result set for every filter.

Transparent concept search—Presents a set of features that enable a more defensible, interactive, and accurate concept search process: Concept search preview—Allows users to contextually refine the concept search by previewing the most frequently occurring terms and selecting only the relevant ones.

Concept search explorer—Presents a visual interface that enables users to dynamically construct searches by exploring terms and linking them to form comprehensive and relevant concepts.

Concept search report—Automatically documents the related terms included in each concept search and provides detailed analytics of search results.

First-Pass Reveiw and Export

Optimized review interface—Presents an HTML review interface that maximizes screen real estate and minimizes mouse clicks to increase review throughput. Dual monitor support, breadcrumb navigation, and thumbnail view allow reviewers to quickly navigate through document sets.

Decision-tree tagging—Allows individual items (for example, emails, loose files, attachments, or embeddings) or sets of documents to be tagged using a multi-layer tree structure. This directs reviewers to key decision points, preventing errors, and minimizing the number of clicks needed to accurately tag a document. Customization capabilities allow administrators to create different tag sets and control tag access.

Load file configurator—Provides customizable load file creation during export. Expanded options allow for multiple metadata formats CSV, DAT, EDRM XML, as well as Concordance Relativity®, and Summation® load file

Learn More

Unlock faster, more efficient discovery with Arctera. Experience the power and simplicity firsthand—try Arctera eDiscovery Platform today.